

M. MISITI

Y. MISITI

G. OPPENHEIM

J-M. POGGI

Méthodes d'ondelettes en statistique : introduction et exemples

Journal de la société statistique de Paris, tome 139, n° 4 (1998), p. 3-29

http://www.numdam.org/item?id=JSFS_1998__139_4_3_0

© Société de statistique de Paris, 1998, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

MÉTHODES D'ONDELETTES EN STATISTIQUE : INTRODUCTION ET EXEMPLES

M. MISITI^{1,2}, Y. MISITI¹,
G. OPPENHEIM^{1,3}, J-M. POGGI^{1,4}

¹ Université Paris XI - Orsay, Laboratoire de Mathématiques, UMR C8628,
«Probabilités, Statistique et Modélisation» bât. 425, 91405 Orsay, France

² Ecole Centrale de Lyon

³ Université de Marne-la-Vallée

⁴ Université René Descartes, Paris V

I. INTRODUCTION

Les ondelettes sont un outil récent de traitement du signal, permettant l'analyse à plusieurs échelles de temps, des propriétés locales de signaux non nécessairement stationnaires, et donnant lieu à de nombreuses applications dans des domaines variés. En particulier, elles sont à la base de nouvelles techniques d'analyse et de synthèse et trouvent de remarquables applications dans des domaines larges tels que la compression et le débruitage (exemples développés plus bas); elles fournissent aussi des solutions particulièrement élégantes à des problèmes plus ponctuels comme, par exemple, la génération de trajectoires de processus Browniens fractionnaires (cf. SELLAN (1995) et ABRY, SELLAN (1996)).

Citons quelques références offrant d'amples introductions au domaine des ondelettes tant du point de vue des mathématiques que du point de vue du traitement du signal : DAUBECHIES (1992), KAHANE, LEMARIÉ (1995), LEMARIÉ (1990), MALLAT (1998), MEYER (1990) et (1993), WIRCKERHAUSER (1994). En outre, on trouvera dans le livre de BURKE-HUBBARD (1996), un remarquable texte sur l'histoire des ondelettes.

Après une introduction générale aux ondelettes, dans cet article à vocation pédagogique qui souhaite être accessible aux non spécialistes, on se concentre sur quelques applications en privilégiant les aspects statistiques. Plus précisément, on traite des exemples concrets de signaux réels pour les signaux unidimensionnels puis de signaux multidimensionnels (images), pour illustrer principalement les trois thèmes suivants :

- l'*analyse* ou comment utiliser la transformée en ondelettes (cf. MALLAT (1989) pour les aspects liés à la multirésolution, COIFMAN, WICKERHAUSER (1992) pour les paquets d'ondelettes et ARNEODO *et al.* (1995) pour

d'intéressantes applications de la transformée continue à des signaux présentant des propriétés d'autosimilarité) pour scruter les données et dégager des pistes pour un stade ultérieur du traitement ;

- le *débruitage* ou encore l'estimation de fonctions en se basant sur les représentations par ondelettes, qui donne lieu à des algorithmes très simples et souvent plus performants et faciles à régler que les méthodes traditionnelles d'estimation fonctionnelle grâce à leur adaptativité. On peut trouver dans DONOHO (1995), DONOHO, JOHNSTONE (1994), DONOHO *et al.* (1995), OGDEN (1997), HÄRDLE *et al.* (1998) quelques-uns des résultats typiques de ce domaine, et se référer à ANTONIADIS, OPPENHEIM (1995) pour saisir l'ampleur de la stimulation exercée par ces idées dans la communauté des statisticiens ;
- la *compression* et en particulier la compression des images (cf. DE VORE *et al.* (1992) et COHEN (1995)) où les ondelettes constituent une méthode très compétitive. Pour illustrer ce dernier point, on renvoie à BRISLAWN (1995) pour une application aux empreintes digitales dans laquelle les techniques basées sur les ondelettes se sont révélées meilleures que les autres méthodes.

Les exemples présentés sont traités grâce à la boîte à outils ondelettes de Matlab (cf. MISITI *et al.* (1996)) qui est un logiciel de traitement de signaux ou d'images par ondelettes et par paquets d'ondelettes. Les figures qui suivent sont toutes des copies d'écran obtenues au moyen de ce logiciel. Notons que de nombreux logiciels sont apparus dès le début des années 90, et particulièrement depuis six ans, dans le champ très actif des ondelettes et de leurs applications. On en trouvera une liste très complète dans un article assez récent de BRUCE *et al.* (1996), séparant bien les logiciels commerciaux de ceux qui sont disponibles gratuitement sur le réseau international.

Notons pour terminer cette introduction que de très nombreux thèmes, à la fois théoriques et appliqués, ne sont pas abordés ici. On trouvera à la fin de cet article, en guise de conclusion, quelques éléments complémentaires à leur sujet.

II. ONDELETTES

II.1. Généralités

Soit ψ une fonction choisie suffisamment régulière et bien localisée. Cette fonction ψ sera appelée ondelette si elle vérifie la condition d'admissibilité dans le domaine fréquentiel suivante :

$$\int_{\mathbb{R}^+} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega = \int_{\mathbb{R}^-} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega = C_+ < +\infty,$$

où $\hat{\psi}$ désigne la transformée de Fourier de ψ ,

condition qui entraîne, en particulier, que l'ondelette est d'intégrale nulle. Cette condition minimale, est souvent renforcée en exigeant que l'ondelette

ait des moments nuls : la fonction ψ a $m + 1$ moments nuls si pour tout $k = 0, \dots, m$

$$\int_{\mathbb{R}} t^k \psi(t) dt = 0.$$

Pour fixer les idées, disons qu'une ondelette oscille pendant un certain temps comme une onde, mais est ensuite localisée grâce à un amortissement. L'oscillation d'une ondelette se mesure par le nombre de moments nuls.

A partir de cette unique fonction ψ , on construit par translation et dilatation, une famille de fonctions qui sont les atomes de base :

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a \in \mathbb{R}_+, b \in \mathbb{R}.$$

Pour une fonction f de carré intégrable, on définit sa transformée en ondelettes continue par la fonction C telle que :

$$C(a, b) = \int_{\mathbb{R}} f(t) \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) dt.$$

Calculer cette fonction C c'est faire l'analyse de f par l'ondelette ψ .

Il convient d'observer que, si f est une fonction constante, ses coefficients d'ondelettes sont nuls. Les coefficients d'ondelettes mesurent les fluctuations à l'échelle a , de la fonction f . La tendance à l'échelle a , contenant les évolutions plus lentes, est essentiellement éliminée dans $C(a, b)$. L'analyse en ondelettes permet une analyse locale de f ainsi que la mise en évidence d'effets d'échelle en comparant les $C(a, b)$ pour différentes valeurs de a . En effet, supposons que ψ soit nulle en dehors de $[-M, +M]$, alors $\psi\left(\frac{t-b}{a}\right)$ est nulle en dehors de $[-Ma + b, Ma + b]$. Par conséquent, la valeur de $C(a, b)$ dépend des valeurs de f dans un voisinage de b de longueur proportionnelle à a .

Notons à cet égard que la situation des ondelettes diffère de celle de l'analyse de Fourier, puisque la valeur de la transformée de Fourier de f en un point dépend des valeurs de f sur \mathbb{R} tout entier. Qualitativement les grandes valeurs de $C(a, b)$ fournissent donc une information sur l'irrégularité locale de f , autour de la position b et à l'échelle a . En ce sens, l'analyse en ondelettes est une analyse des fluctuations de f à toutes les échelles. On trouvera dans MALLAT (1998) ou dans ABRY (1998) des compléments quant à la quantification de la notion de localisation et à la comparaison entre les analyses de Fourier et par ondelettes.

La transformée continue (cf. TORRÉSANI, 1995) définie ci-dessus permet de caractériser la régularité höldérienne de fonctions et son utilisation statistique pour la détection de phénomènes transients ou de ruptures est féconde.

Dans de nombreuses situations (et partout dans la suite de ce texte), on se limite aux valeurs de a et de b suivantes :

$$a = 2^j, \quad b = k2^j = ka \quad \text{pour } (j, k) \in \mathbb{Z}^2.$$

Dans ce cas et pour des ondelettes vérifiant des propriétés plus fortes, en particulier le cas orthogonal (précisé plus bas) dans lequel nous nous placerons désormais, une fonction, appelée fonction d'échelle et notée ϕ , est associée à ψ . Elle est dilatée et translatée comme ψ . En deux mots, la fonction ϕ est aux approximations locales ce qu'est la fonction ψ pour les fluctuations autour de la tendance locale.

On définit alors les atomes de base des ondelettes :

$$\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k), \quad \text{pour } (j, k) \in \mathbb{Z}^2,$$

et

$$\phi_{j,k}(x) = 2^{-j/2}\phi(2^{-j}x - k), \quad \text{pour } (j, k) \in \mathbb{Z}^2.$$

Dans ce contexte les coefficients d'ondelettes du signal s sont donnés par

$$\alpha_{j,k} = \int_{\mathbb{R}} s(t)\psi_{j,k}(t)dt$$

et, sous certaines conditions (vérifiées pour une ondelette orthogonale), ces coefficients suffisent pour reconstruire le signal par

$$s(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \alpha_{j,k} \psi_{j,k}(t).$$

L'existence d'une fonction ψ telle que la famille $(\psi_{j,k})_{(j,k) \in \mathbb{Z}^2}$ soit une base orthonormée de $L^2(\mathbb{R})$ est intimement liée à la notion d'analyse multirésolution (cf. MEYER 1990 et MALLAT 1989, 1998) notée AMR en abrégé. Une AMR de $L^2(\mathbb{R})$ consiste en une suite de sous-espaces fermés emboîtés $(V_j)_{j \in \mathbb{Z}}$:

$$\dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \dots$$

de $L^2(\mathbb{R})$, dont l'intersection est triviale et l'union dense dans $L^2(\mathbb{R})$.

Ces espaces sont tous déduits de l'espace «central» V_0 par dilatation (pour $j < 0$) ou contraction (pour $j > 0$), c'est-à-dire :

$$f(t) \in V_j \Leftrightarrow f(2t) \in V_{j-1} \quad \text{pour } j \in \mathbb{Z}.$$

Enfin, il existe une fonction ϕ de V_0 telle que

$$V_0 = \{f \in L^2(\mathbb{R}) \mid f(t) = \sum_{k \in \mathbb{Z}} e_k \phi(t - k)\}.$$

La fonction ϕ est celle introduite ci-dessus.

Les $(V_j)_{j \in \mathbb{Z}}$ de l'AMR servent d'espaces d'approximation (ou de tendance). On définit aussi les espaces de détail notés $(W_j)_{j \in \mathbb{Z}}$ et définis par :

$$V_j = V_{j+1} \oplus W_{j+1} \quad \text{pour } j \in \mathbb{Z}.$$

Les $(\phi_{j,k})_{k \in \mathbb{Z}}$ engendrent V_j tandis que les $(\psi_{j,k})_{k \in \mathbb{Z}}$ engendrent W_j . Comme $L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j$, les $(\psi_{j,k})_{(j,k) \in \mathbb{Z}^2}$ forment une base orthonormée d'ondelettes de $L^2(\mathbb{R})$. Ainsi $\alpha_{j,k}$ est le coefficient associé à $\psi_{j,k}$ dans la projection orthogonale de s sur W_j .

Les rôles respectifs des fonctions ϕ et ψ ainsi que les notions de détail et de tendance seront illustrés plus bas sur les exemples.

II.2. Une ondelette

La figure 1 concerne l'ondelette notée db4 due à I. DAUBECHIES. En haut, de gauche à droite, on trouve la fonction d'échelle puis l'ondelette. Comme on le voit facilement, l'ondelette ψ oscille et est d'intégrale nulle tandis que la fonction d'échelle ϕ oscille moins et est d'intégrale positive (en fait égale à 1). Calculer le produit scalaire d'une fonction ou d'un signal avec l'ondelette permet d'analyser les fluctuations du signal autour d'une moyenne locale fournie par le calcul du produit scalaire du signal avec la fonction d'échelle.

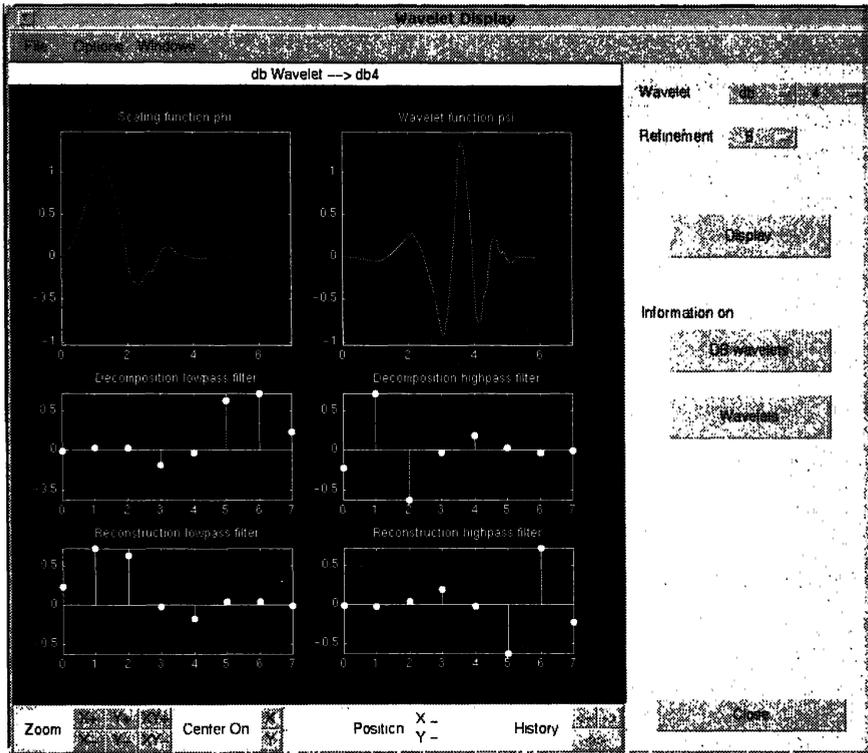


FIGURE 1 : L'ondelette notée db4 due à I. Daubechies

Plus bas on trouve les quatre filtres utiles pour effectuer les calculs en utilisant un algorithme rapide dû à S. MALLAT. Ces filtres vont par paire. Deux sont associés à la fonction d'échelle ϕ et figurent dans la première colonne, et les deux autres filtres, dans la seconde colonne, sont associés à l'ondelette ψ .

Il est facile de voir que dans une colonne on passe de l'un à l'autre en prenant le filtre miroir et que dans une ligne on passe de l'un à l'autre en prenant le filtre miroir puis en multipliant les termes d'indice pair par -1 . Ainsi, un seul de ces filtres suffit à fabriquer tous les autres. De même, on peut montrer qu'à partir de la donnée d'une AMR et donc de la fonction d'échelle, on peut construire l'ondelette. Ces filtres apparaissent dans les relations qu'entretiennent les fonctions de base associées à des niveaux successifs.

Ce sont les équations aux deux échelles suivantes :

$$\phi_{j+1,0} = \sum_{k \in \mathbb{Z}} h_k \phi_{j,k}$$

et

$$\psi_{j+1,0} = \sum_{k \in \mathbb{Z}} g_k \phi_{j,k}.$$

La suite h détermine les filtres de la première colonne et la suite g détermine ceux de la deuxième colonne.

L'ondelette présentée figure 1 fait partie d'une famille d'ondelettes indexée par $n \in \mathbb{N}^*$ et due à I. DAUBECHIES, 1990 (cf. p. 194-202 pour la construction). L'ondelette *db1* est simplement l'ondelette de Haar. Les principales propriétés de l'ondelette *dbn* sont :

- c'est une ondelette orthogonale, associée à une AMR;
- elle est à support compact $[0, 2n - 1]$ et les filtres associés sont de longueur $2n$;
- le nombre de moments nuls est n et elle est en général très asymétrique;
- la régularité est $0.2n$ lorsque n est suffisamment grand.

II.3. Organisation des ondelettes

Les ondelettes sont donc organisées en utilisant deux paramètres :

- le temps k , permettant de translater les formes pour un niveau donné;
- l'échelle 2^j , permettant de passer d'un niveau j au niveau immédiatement inférieur dans l'arbre sous-jacent.

On trouve dans la première colonne de la figure 2, les dilatées dyadiques (2 fois, 4 fois, 8 fois,...) de la fonction d'échelle et dans la seconde colonne, celles de l'ondelette. Les fonctions de la première colonne sont associées au calcul des moyennes locales du signal (on dira approximations) et sont de la forme :

$$A_j(t) = \sum_{k \in \mathbb{Z}} \beta_{j,k} \phi_{j,k}(t)$$

avec, ce que nous appellerons les coefficients d'approximation :

$$\beta_{j,k} = \int_{\mathbb{R}} s(t) \phi_{j,k}(t) dt,$$

tandis que celles de la seconde colonne sont associées au calcul des différences entre deux moyennes locales successives, ce sont des retouches (on dira détails) et sont de la forme

$$D_j(t) = \sum_{k \in \mathbb{Z}} \alpha_{j,k} \psi_{j,k}(t).$$

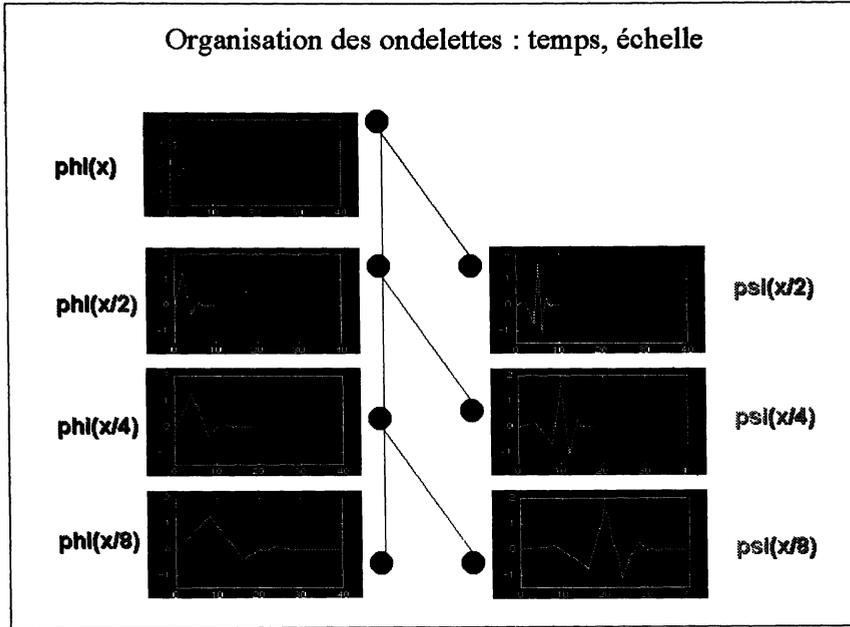


FIGURE 2 : Organisation des ondelettes

En résumé, on dispose donc de quatre sortes d'objets :

- les coefficients d'ondelettes : $\alpha_{j,k} = \int_{\mathbb{R}} s(t)\psi_{j,k}(t)dt$ permettant de définir les détails ;
- les signaux de détail eux-mêmes : $D_j(t) = \sum_{k \in \mathbb{Z}} \alpha_{j,k}\psi_{j,k}(t)$;
- les coefficients d'approximation : $\beta_{j,k} = \int_{\mathbb{R}} s(t)\phi_{j,k}(t)dt$ permettant de calculer les approximations directement sans passer par des sommes de détails ;
- les signaux d'approximation eux-mêmes : $A_j(t) = \sum_{k \in \mathbb{Z}} \beta_{j,k}\phi_{j,k}(t)$.

Les signaux de détail et d'approximation sont des fonctions de t , le temps du signal d'origine, alors que les coefficients de niveau j sont dans le temps dyadique $2^j \mathbb{Z}$. La notion d'AMR introduite précédemment permet d'interpréter les notions de détail et de tendance en termes de projection orthogonale sur les espaces W_j et V_j respectivement. En d'autres termes, si l'on assimile le signal discret s à des coefficients $(\beta_{0,k})_{k \in \mathbb{Z}}$ dans V_0 , les $(\alpha_{j,k})_{k \in \mathbb{Z}}$ et les $(\beta_{j,k})_{k \in \mathbb{Z}}$ s'interprètent comme les coefficients de s dans les bases de W_j et de V_j respectivement tandis que D_j et A_j sont des éléments des espaces W_j et V_j repérés comme des fonctions de V_0 .

II.4. L'arbre d'ondelettes pour un signal

Un tel arbre est présenté dans la figure 3. On trouve à sa racine un signal S (on peut dire aussi une chronique). L'arbre peut se lire de diverses manières. La première colonne donne trois approximations, de la plus fine A_1 à la plus grossière A_3 , comme on peut facilement s'en convaincre en se focalisant sur la fin du signal. Les différences entre deux approximations successives sont capturées dans les détails notés D_1 à D_3 , plus précisément on a $D_1 = S - A_1$, $D_2 = A_1 - A_2$, et donc $S = A_2 + D_2 + D_1$.

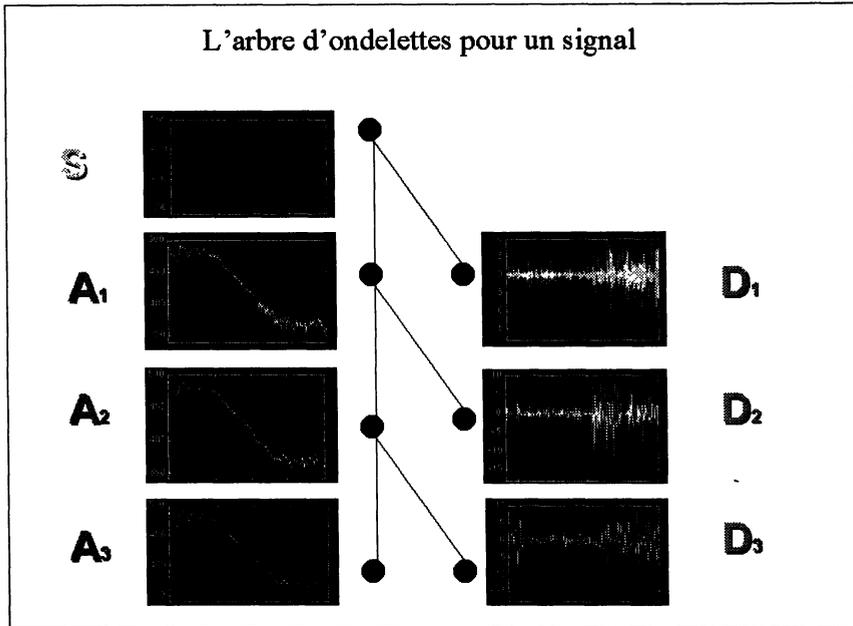


FIGURE 3 : L'arbre d'ondelettes pour un signal

Plus généralement, si l'on reprend «l'égalité» signifiant que l'on peut reconstruire le signal à partir de ses coefficients $\alpha_{j,k}$, c'est-à-dire

$$s(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \alpha_{j,k} \psi_{j,k}(t),$$

on peut l'utiliser pour définir le détail au niveau j . Fixons j et sommions sur k . Un détail D_j n'est rien d'autre que

$$D_j(t) = \sum_{k \in \mathbb{Z}} \alpha_{j,k} \psi_{j,k}(t),$$

puis sommions sur j : on retrouve le fait que le signal est la somme de ses détails

$$s = \sum_{j \in \mathbb{Z}} D_j.$$

Les détails sont définis. Prenons maintenant un niveau de référence noté J (3 dans le présent exemple). Il y a deux sortes de détails. Ceux associés aux indices $j \leq J$, correspondant aux échelles $a = 2^j \leq 2^J$, qui sont les détails plus fins que la résolution correspondant à J . Les autres, pour lesquels $j > J$, sont plus grossiers. Groupons ces derniers :

$$A_J = \sum_{j>J} D_j,$$

cela définit ce que nous appellerons l'approximation au niveau J du signal s . On a la relation

$$s = A_J + \sum_{j \leq J} D_j,$$

signifiant que s est la somme de son approximation A_J et des détails plus fins. On en déduit que les approximations sont liées par

$$A_{J-1} = A_J + D_J.$$

Dans le cas orthogonal, pour lequel $(\psi_{j,k})_{j,k}$ est une famille orthogonale, on a :

- A_J est orthogonal à $D_J, D_{J-1}, D_{J-2}, \dots$,
- s est la somme de deux signaux orthogonaux : A_J et $\sum_{j \leq J} D_j$,
- la qualité Q_J de l'approximation de s par A_J est égale à $Q_J = \frac{\|A_J\|^2}{\|s\|^2}$, et
on a $Q_{J-1} = Q_J + \frac{\|D_J\|^2}{\|s\|^2}$ (cette mesure de la qualité est dite en «énergie»).

III. UN SIGNAL DE CONSOMMATION ÉLECTRIQUE ANALYSÉ PAR ONDELETTES

Comme premier exemple, nous considérons une chronique de la consommation électrique minute par minute de la France métropolitaine; le problème est présenté en détail dans MISITI *et al.* (1994).

Dans la figure 4, on trouve de haut en bas, on trouve le signal d'origine, l'approximation au niveau 5 et les détails du niveau le plus grossier ($d5$) au niveau le plus fin ($d1$). Tous les signaux présentés sont dans la même unité de temps, ce qui permet en particulier une lecture synchrone en temps de tous les graphiques. L'ondelette utilisée est db3.

Le signal analysé représente, à une transformation près, la consommation électrique de trois jours, en période estivale en France métropolitaine. Les trois jours sont un jeudi suivi d'un vendredi, qui sont de forme et de niveau très semblables, puis un samedi qui est de niveau moyen beaucoup plus bas à cause du début de week-end et de la moindre activité économique. Le pas de temps est la minute. On peut ainsi réinterpréter les échelles de temps correspondant

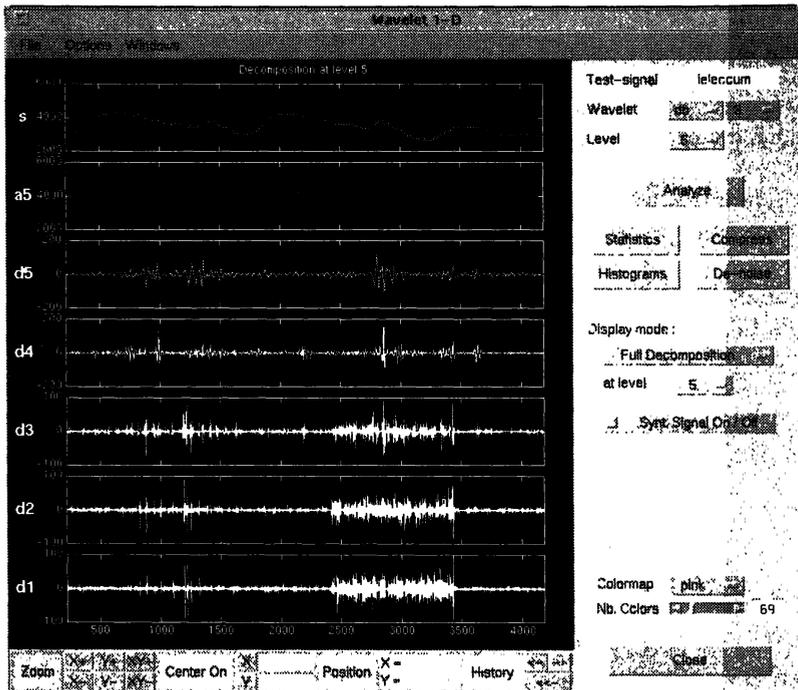


FIGURE 4 : Un signal de consommation électrique analysé par ondelettes

à chacun des niveaux. En négligeant les effets du choix de l'ondelette, on peut dire qu'approximativement $d1$ contient les composantes du signal de période entre 1 et 2 minutes, $d2$ celles de période entre 2 et 4 minutes, ainsi de suite jusqu'à $d5$ qui contient les composantes du signal de période entre 16 et 32 minutes. L'approximation $a5$ contient les composantes du signal de période supérieure à 32 minutes.

Un examen rapide montre que :

- l'analyse permet de repérer des points aberrants possibles, ils se lisent grâce aux très grandes valeurs de $d1$ autour de l'abscisse 1200 ;
- on distingue facilement dans les graphiques des détails $d1$ et $d2$, les bruits de mesure et d'état, dont l'amplitude est faible en régime normal, qui donnent ces détails oscillant rapidement autour de 0 ;
- par opposition, on isole facilement la période de dysfonctionnement des capteurs correspondant à la longue séquence de valeurs anormalement grandes de $d1$, $d2$ et plus faiblement $d3$, entre l'abscisse 2500 et l'abscisse 3500 ;
- on peut remarquer en outre que, dans les détails $d4$ et $d5$, on ne distingue plus cette période, la différence due à un bruit exceptionnel des capteurs s'évanouit, ce qui donne une indication sur le contenu fréquentiel de ce bruit additionnel ;

- la pseudo-périodicité hebdomadaire manifeste dans la trajectoire analysée peut se lire, pour les échelles examinées, dès les niveaux 4 et 5 par l'apparition de motifs périodiques dans les détails, à des instants périodiques. On peut donc en déduire que les détails d_4 et d_5 contiennent des composantes du signal utile par opposition au bruit non informatif. Ce phénomène ne se produisant pas pour les détails de niveau plus fin, on peut penser qu'ils sont presque intégralement constitués de bruit.

Cet exemple rapide illustre qu'une simple analyse par ondelettes peut donner des pistes nombreuses pour des traitements plus fins et orienter la stratégie de résolution de problèmes classiques comme la détection d'observations aberrantes, le débruitage ou l'extraction du signal utile.

IV. DÉBRUITAGE PAR ONDELETTES : AVANT ET APRÈS

Le débruitage est l'application majeure des ondelettes en statistique. Ce problème, dont un exemple de résolution possible est esquissé dans la page précédente, admet une solution très élégante.

On voit dans la figure 5 un signal bruité réel (en haut) qui est une portion de celui analysé dans les planches précédentes.

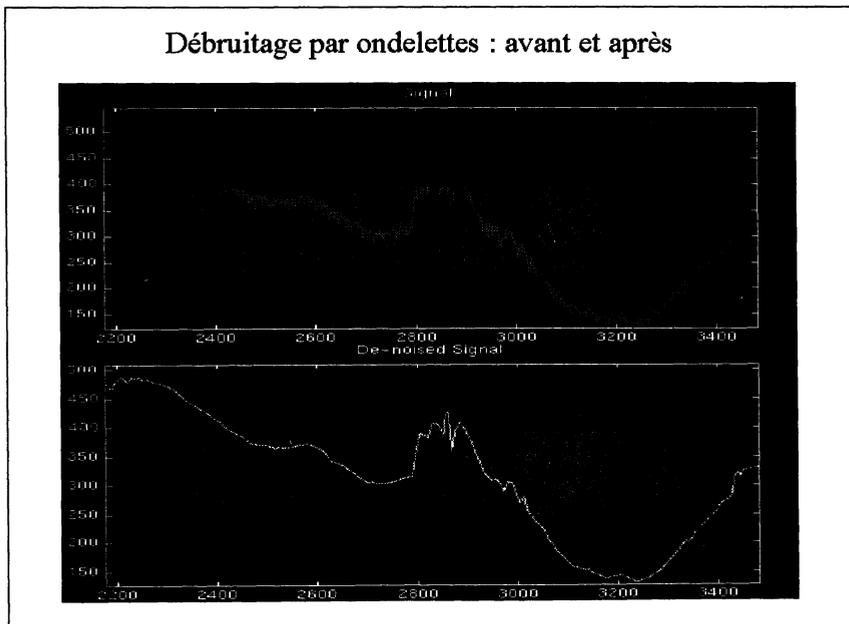


FIGURE 5 : Débruitage par ondelettes : avant et après

MÉTHODES D'ONDELETTES EN STATISTIQUE

Le signal débruité par ondelettes est situé en dessous. Il est manifestement bien débruité, tant dans les zones où le signal est lisse que dans la zone autour de la position 2900 où le signal présente des irrégularités. Les méthodes classiques sont incapables d'une telle adaptation en temps.

Ces méthodes classiques d'estimation de la fonction f dans le modèle

$$Y_i = f(t_i) + \epsilon_i \quad t_i = i/n \quad i = 1, \dots, n$$

où f est inconnue, les $(Y_i)_i$ sont observés et ϵ est un bruit blanc gaussien inobservable, sont essentiellement de deux sortes. On trouve, d'une part, les méthodes de noyau conduisant à estimer $f(t)$ par une moyenne locale des valeurs de Y_i pour des t_i proches de t en utilisant un noyau K régularisant

$$\hat{f}(t) = \sum_i \frac{K(t - t_i)}{\sum_i K(t - t_i)} Y_i$$

et d'autre part, les méthodes par projection orthogonale qui supposent que la fonction f appartient à un espace \mathcal{F} pour lequel on dispose d'une famille orthogonale $(e_k)_k$, de sorte que f s'écrive comme

$$f = \sum_k \gamma_k e_k$$

L'estimateur de f s'en déduit par estimation des γ_k à l'aide des Y_i

$$\hat{f} = \sum_k \hat{\gamma}_k e_k$$

Les premiers travaux utilisant les ondelettes proposaient pour des fonctions f de carré intégrable, des estimateurs de la forme

$$\hat{f} = \sum_k \hat{\beta}_{J,k} \phi_{J,k}$$

qui sont, à la fois des estimateurs à noyau (en prenant ϕ pour K) et par projection orthogonale (en prenant les $\phi_{j,k}$ pour les e_i), et qui consistent à estimer l'approximation de f au niveau J . Les estimateurs par ondelettes postérieurs complètent cette estimation par des termes de détail

$$\hat{f} = \sum_k \hat{\beta}_{J,k} \phi_{J,k} + \sum_{J \leq j \leq J_0} \sum_k \hat{\alpha}_{j,k} \psi_{j,k}$$

en choisissant astucieusement les $\hat{\alpha}_{j,k}$. Cela permet l'adaptativité désirée.

Voici comment.

V. ANALYSE PAR ONDELETTES D'UN SIGNAL DE TYPE DOPPLER

Considérons tout d'abord, un signal de synthèse permettant de bien saisir l'esprit de la technique de débruitage par ondelettes (figure 6).

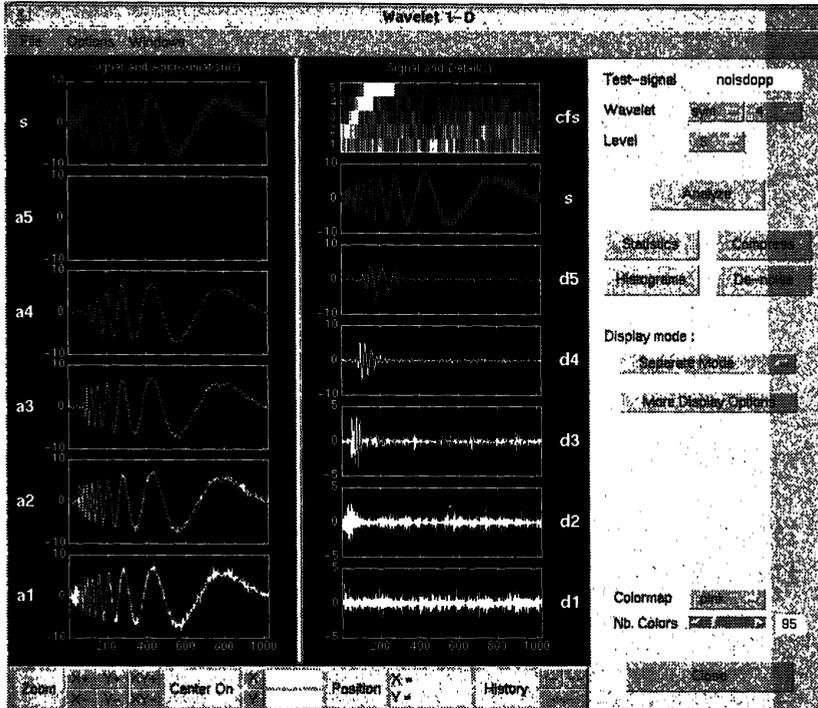


FIGURE 6 : Analyse par ondelettes d'un signal de type Doppler

L'écran est organisé en deux colonnes. Dans la première colonne, on voit le signal bruité s puis en dessous les approximations du niveau 5 (le plus grossier, $d5$) au niveau 1 (le plus fin $d1$).

Dans la seconde colonne, on voit en haut une version coloriée des coefficients d'ondelettes des niveaux 1 à 5 (cfs), suivie du signal bruité puis en dessous les détails du niveau 5 (le plus grossier, $d5$) au niveau 1 (le plus fin $d1$).

L'ondelette utilisée est une ondelette à support compact presque symétrique d'ordre 4 notée $sym4$.

Commençons par examiner la première colonne et concentrons-nous sur la portion du signal correspondant aux abscisses 200 à 1000. Partons de $a1$ et recherchons, en remontant, un niveau tel que l'approximation constitue un bon candidat à être un estimateur du signal utile. Un niveau raisonnable est le niveau 4 (ou le niveau 5). Néanmoins l'estimateur associé à $a4$ est clairement très mauvais pour le début du signal correspondant aux abscisses

0 à 100. Inversement une restitution acceptable du début du signal conduirait à choisir a_2 visiblement trop bruité.

Regardons maintenant les détails. Le détail d_1 semble composé intégralement de bruit. Les détails d_2 à d_5 présentent de fortes valeurs concentrées aux abscisses de 0 à 100. Ceci est aussi visible sur le graphique des coefficients d'ondelettes, les plus grands en valeur absolue étant les plus clairs. Cette forme vient du fait que le signal est une fonction sinusoïdale dont l'amplitude croît avec le temps : les oscillations aux échelles les plus petites expliquent les détails montrés, les autres sont dans l'approximation a_5 (que l'on peut simplement deviner sur le graphique).

Il en résulte qu'une stratégie plausible de débruitage consiste à :

- conserver une approximation telle que le bruit soit absent ou très atténué (a_4 ou a_5);
- compléter cette approximation par les parties des détails plus fins clairement attribuables au signal utile, en rejetant les parties que l'on considère comme provenant du bruit.

C'est précisément ce que font les méthodes de débruitage par ondelettes mais de façon automatique. Le choix *ad-hoc* effectué ci-contre en connaissant le signal est celui qui est fait par l'une des méthodes les plus répandues de débruitage par ondelettes due à Donoho et Johnstone.

VI. DÉBRUITAGE PAR ONDELETTES D'UN SIGNAL DE TYPE DOPPLER

Considérons la figure 7. L'écran est organisé en deux colonnes. Dans la première colonne, on voit les coefficients d'ondelettes du niveau 5 au niveau 1. Pour les rendre plus lisibles, ils sont «bégayés» 2^k fois au niveau k (ce qui explique les créneaux surtout visibles pour $k > 3$). Dans chacun de ces graphiques, on note la présence de deux droites horizontales en pointillés : les coefficients à l'intérieur du tube sont mis à zéro par la procédure de débruitage. Dans la seconde colonne, le signal bruité s est superposé au signal débruité. En dessous on trouve une version coloriée des coefficients d'ondelettes des niveaux 1 à 5 du signal bruité d'origine et dans le graphique d'en dessous l'homologue pour les coefficients d'ondelettes seuillés, à partir desquels est reconstitué le signal débruité.

L'ondelette utilisée est $\text{sym}4$ comme plus haut.

Voyons, sur cet exemple, comment fonctionne le débruitage :

- les coefficients de l'approximation sont intégralement conservés, ce qui conduit d'emblée à introduire le premier composant du signal débruité qui est de la forme (\hat{A}_d n'est pas représenté directement sur cette planche)

$$\hat{A}_d = \sum_k \hat{\beta}_{5,k} \phi_{5,k}$$

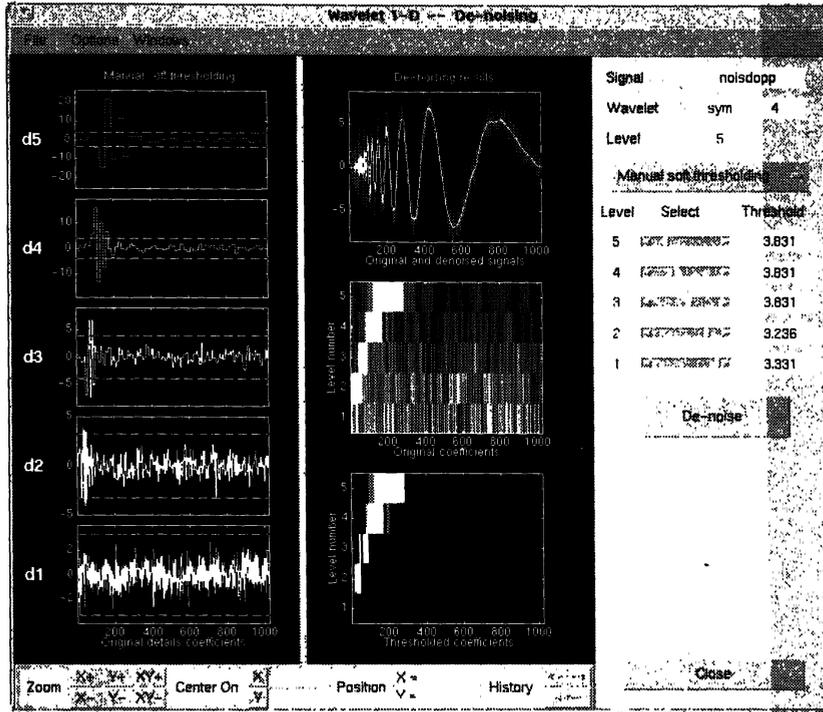


FIGURE 7 : Débruitage par ondelettes d'un signal de type Doppler

Les coefficients estimés $\hat{\beta}_{5,k}$ sont simplement les coefficients obtenus par décomposition du signal initial;

- seuls les quelques plus grands coefficients d'ondelettes (c'est-à-dire les coefficients de détail) sont conservés, les autres sont remplacés par zéro. On peut regarder dans la deuxième colonne, les coefficients avant et après cette opération de seuillage, pour constater la rareté des coefficients conservés. On obtient ainsi le second composant du signal débruité qui est de la forme

$$\hat{D}_d = \sum_{1 \leq j \leq 5} \sum_k \hat{\alpha}_{j,k} \psi_{j,k}$$

Les coefficients estimés $\hat{\alpha}_{j,k}$ sont simplement les coefficients $\alpha_{j,k}$ obtenus par décomposition du signal initial puis seuillés.

De nombreuses méthodes automatiques sont disponibles pour le choix des seuils, suivant la forme et les hypothèses concernant le modèle censé représenter convenablement la manière d'engendrer les données. Nous ne les détaillerons pas ici;

- le signal débruité est donc

$$\hat{s}_d = \hat{A}_d + \hat{D}_d$$

Il est visible que le résultat obtenu en haut à droite, est de bonne qualité, sauf pour le tout début du signal, qui vibre trop à petite échelle avec des fluctuations petites vis à vis de celles du bruit.

VII. DÉBRUITAGE PAR ONDELETTES D'UN SIGNAL DE CONSOMMATION ÉLECTRIQUE

Le premier paragraphe de cette section peut être omis lors d'une lecture séquentielle de l'article.

L'écran présenté dans la figure 8, est organisé en deux colonnes. Dans la première colonne, on voit les coefficients d'ondelettes du niveau 5 au niveau 1, ils sont «bégayés» 2^k fois au niveau k (ce qui explique les créneaux surtout visibles pour $k > 3$). Dans chacun de ces graphiques, on note la présence de deux droites horizontales en pontillés : les coefficients à l'intérieur du tube sont mis à zéro par la procédure de débruitage. Dans la seconde colonne, on voit le signal bruité s superposé au signal débruité. En dessous, on trouve une version colorisée des coefficients d'ondelettes des niveaux 1 à 5 du signal bruité d'origine et dans le graphique d'en dessous l'homologue pour les coefficients d'ondelettes seuillés, à partir desquels est reconstitué le signal débruité. L'ondelette utilisée est coif5 .

Revenons au problème de débruitage d'un signal réel, qui est plus difficile puisque le bruit est de nature inconnue. La méthode employée ici s'inspire de la situation précédente en adaptant le seuillage au niveau d'échelle de sorte que l'on n'exige plus qu'une homogénéité du bruit par niveau d'échelle.

On voit en examinant la première colonne, que les coefficients des niveaux 1 à 3 ont été tous considérés par la méthode comme attribuables au bruit et qu'ils sont tous mis à zéro. Ceci est conforme aux conclusions de l'analyse de la courbe de charge électrique sur trois jours effectuée auparavant. En revanche pour les niveaux 4 et 5, la procédure sélectionne des coefficients d'ondelettes essentiellement dans la zone autour de l'abscisse 400, permettant ainsi une excellente restitution des fluctuations abruptes du signal.

On retrouve ici une des clés pour la compréhension de l'effectivité de ces méthodes. Citons quelques arguments :

- la décomposition par ondelettes est une analyse additive, par conséquent, l'analyse de $Y_t = f(t) + \epsilon_t$ est égale à la somme des analyses du signal $f(t)$ et du bruit ϵ_t ;
- si l'on suppose que le bruit ϵ est blanc, alors les coefficients d'ondelettes à toutes les échelles sont des bruits blancs de même variance.

D'autre part, les signaux utiles sont, dans bien des cas, assez réguliers sauf en de rares endroits (début et fin de phénomènes transitoires, ruptures par exemple), ce qui rend la décomposition par ondelettes de $f(t)$ très creuse, et très bien représentée par les coefficients d'une approximation assez grossière plus quelques grands coefficients de détail ;

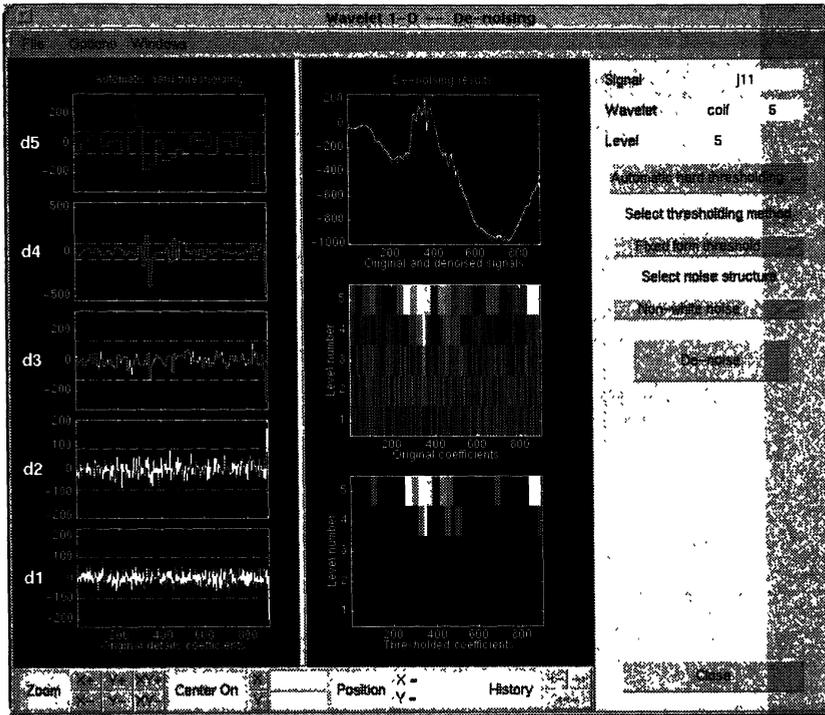


FIGURE 8 : Débruitage par ondelettes d'un signal de consommation électrique

- dans le cas où les irrégularités engendrent des coefficients plus grands que l'échelle du bruit, la procédure de seuillage sélectionne seulement les coefficients liés au signal pourvu que l'on sache estimer convenablement l'échelle du bruit. Enfin l'opération de seuillage conduit toujours à régulariser le signal.

VIII. UNE IMAGE («BARBARA») DÉCOMPOSÉE PAR ONDELETTES

VIII.1. Décomposition sous forme d'arbre

Passons maintenant au traitement d'images par ondelettes et examinons d'abord la figure 9. Cette planche est difficile à lire mais sa structure est informative; les détails se verront mieux dans la planche suivante. On trouve en haut à gauche l'image originale, puis en dessous un tableau d'images à trois lignes et quatre colonnes. L'ensemble constitue l'arbre de la décomposition par ondelettes en 2D, c'est un arbre quaternaire alors qu'en 1D c'est un arbre binaire.

En 1D, le signal est divisé en deux : une approximation et un détail; en 2D, l'image est décomposée en quatre : une approximation (première colonne du

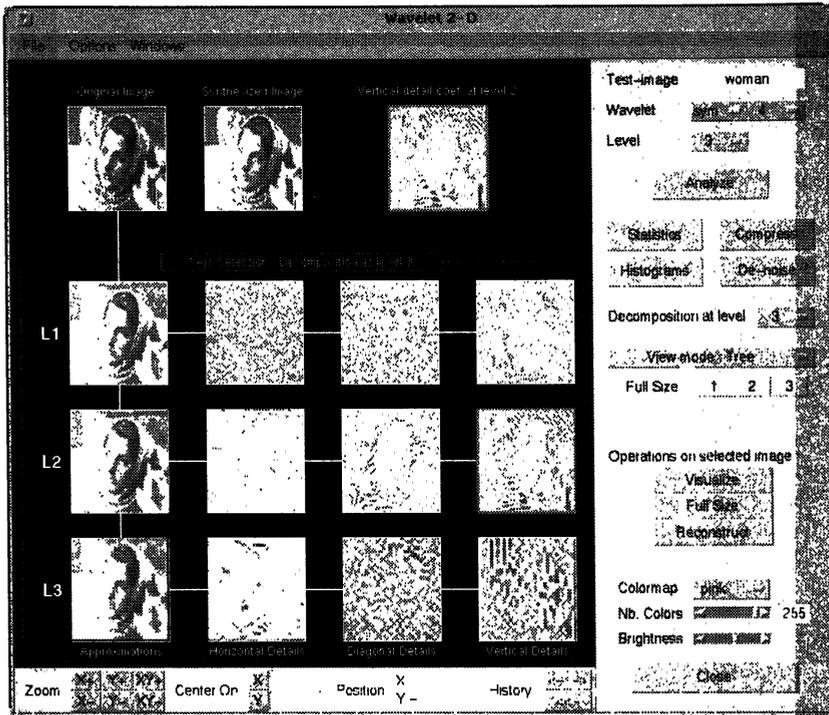


FIGURE 9 : Une image («Barbara») décomposée par ondelettes.
Décomposition sous forme d'arbre

tableau 3×4) et trois détails dans trois directions, horizontale, diagonale et verticale (les trois dernières colonnes du tableau 3×4). Les lignes de ce tableau sont indexées par les niveaux : du niveau 1 (le plus fin) au niveau 3 (le plus grossier).

L'ondelette utilisée ici est sym4.

Avant de commenter plus avant cet écran, disons quelques mots sur les ondelettes orthogonales en 2D. Il s'agit ici de cas très particuliers d'ondelettes, qui sont, et de très loin, les plus utilisées car elles conduisent à des calculs rapides, ce qui est critique en traitement d'images. En 1D, on dispose des deux fonctions ϕ et ψ et des deux filtres associés. A partir de ces objets, on fabrique les quatre ondelettes suivantes qui agissent, non plus sur la droite réelle, mais sur le plan :

- $\phi(x, y) = \phi(x)\phi(y)$ qui sert à définir les approximations,
- $\psi_1(x, y) = \psi(x)\phi(y)$ qui sert à définir les détails horizontaux,
- $\psi_2(x, y) = \psi(x)\psi(y)$ qui sert à définir les détails diagonaux,
- $\psi_3(x, y) = \phi(x)\psi(y)$ qui sert à définir les détails verticaux.

La simplicité algorithmique provient du fait qu'il s'agit simplement de faire opérer successivement sur les lignes puis les colonnes de la matrice associée à

MÉTHODES D'ONDELETTES EN STATISTIQUE

l'image, les deux filtres (à basse et haute fréquence respectivement) utiles en 1D.

En examinant soigneusement les diverses approximations, on constate qu'il s'agit de versions de plus en plus grossières de l'image originale.

En comparant l'image originale avec les coefficients de l'approximation au niveau 3, il est clair que des détails à petite échelle ont été perdus comme par exemple le motif du châle ou la trame du fauteuil en rotin derrière le visage. Les coefficients des détails sont plus difficiles à distinguer. Néanmoins on peut deviner des coefficients forts dans les zones mentionnées ci-contre, surtout aux niveaux 2 et 3, et dans les directions diagonale et verticale, mais aussi des détails du visage comme on peut le voir au niveau 3.

VIII.2. Décomposition sous forme compacte

La disposition précédente de la décomposition insiste sur la structure arborescente; celle qui est proposée ici, insiste sur le codage de la décomposition 2D.

En effet, on trouve figure 10 en haut à gauche, l'image originale et en bas à droite les coefficients de la décomposition de cette image.

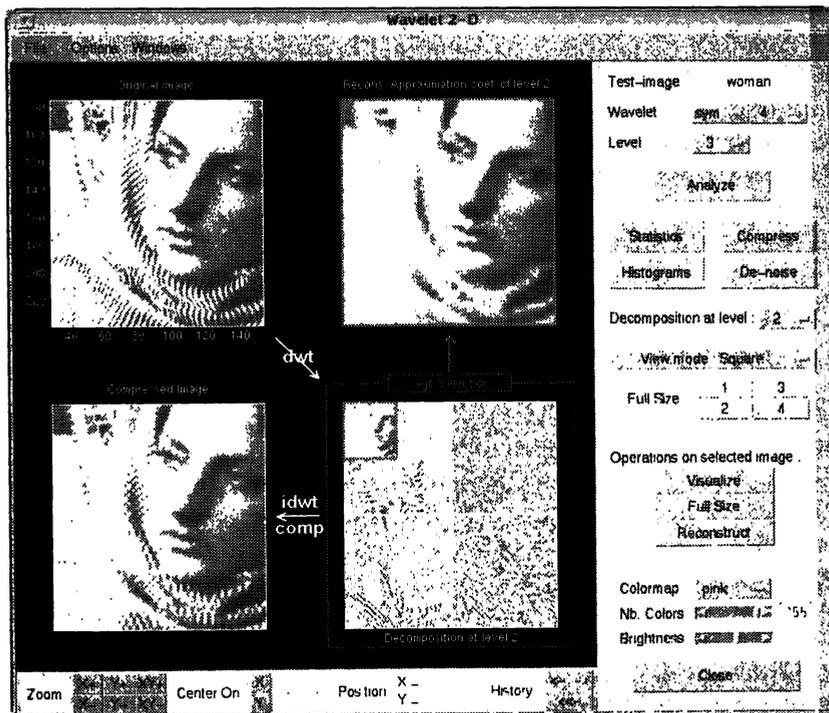


FIGURE 10 : Une image («Barbara») décomposée par ondelettes. Décomposition sous forme compacte

Pour lire ce graphique, il faut partir de l'angle situé en bas à droite : les trois premières imasettes rencontrées représentent (dans le sens inverse des aiguilles d'une montre) les coefficients des détails horizontal, diagonal et vertical au niveau le plus fin (niveau 1), les trois imasettes suivantes représentent les coefficients des détails au niveau 2, et ainsi de suite jusqu'à ce que l'on rencontre une imasette unique (c'est le cas en haut à gauche) qui contient les coefficients de l'approximation au niveau le plus grossier (ici 2).

Les deux autres images sont en bas à gauche, l'image issue d'une procédure de compression (qui sera présentée dans la planche suivante) et en haut à droite, l'image reconstituée à partir des seuls coefficients de l'approximation de niveau 2.

L'ondelette utilisée est encore sym4.

L'image analysée ici est un zoom de la précédente et permet, de ce point de vue, des commentaires plus fins. Concentrons-nous sur les images figurant sur l'anti-diagonale de l'écran, elles sont dans la même résolution que l'image originale, on peut donc les comparer directement.

Examinons l'image située en haut à droite, en la comparant à l'image originale. On retrouve les éléments mentionnés lors du commentaire de la planche précédente. Le motif du châte a été perdu dans l'approximation de niveau 2, ces motifs n'étant discernables qu'à des échelles inférieures à 4 pixels. On ne récupère ici qu'une texture assez homogène avec les deux plis principaux bien restitués et le pli intermédiaire légèrement moins bien défini. Il en est de même du fauteuil qui est difficile à reconnaître. En outre, les traits du visage sont plus flous.

Cette perte générale de définition des contours peut être grandement améliorée en rajoutant de façon sélective des coefficients de détails convenablement choisis. En effectuant une telle opération, on obtient l'image comprimée située en bas à gauche. Elle n'est pas parfaite mais néanmoins beaucoup plus fidèle.

Voici comment elle a été obtenue.

IX. UNE IMAGE («BARBARA») COMPRIMÉE PAR ONDELETTES

On trouve en haut à gauche dans la figure 11, l'image originale qui est à compresser et à côté l'image comprimée.

En dessous, pour chacun des niveaux du plus fin (niveau 1) au plus grossier (ici 3) on trouve pour chacune des orientations, horizontale, diagonale et verticale, l'histogramme des niveaux de gris des coefficients du détail correspondant. Dans chacun de ces histogrammes, on note la présence de deux droites verticales en pointillés : les coefficients à l'intérieur du tube sont mis à zéro par la procédure de compression.

Le pourcentage de zéros de la représentation par ondelettes, qui est un indicateur de la place occupée par l'image comprimée, est de 92 %. Il s'agit d'une méthode de compression avec perte : 97 % de l'énergie (qui n'est pas un indicateur très parlant et pertinent pour les images) est conservée. Pour

MÉTHODES D'ONDELETTES EN STATISTIQUE

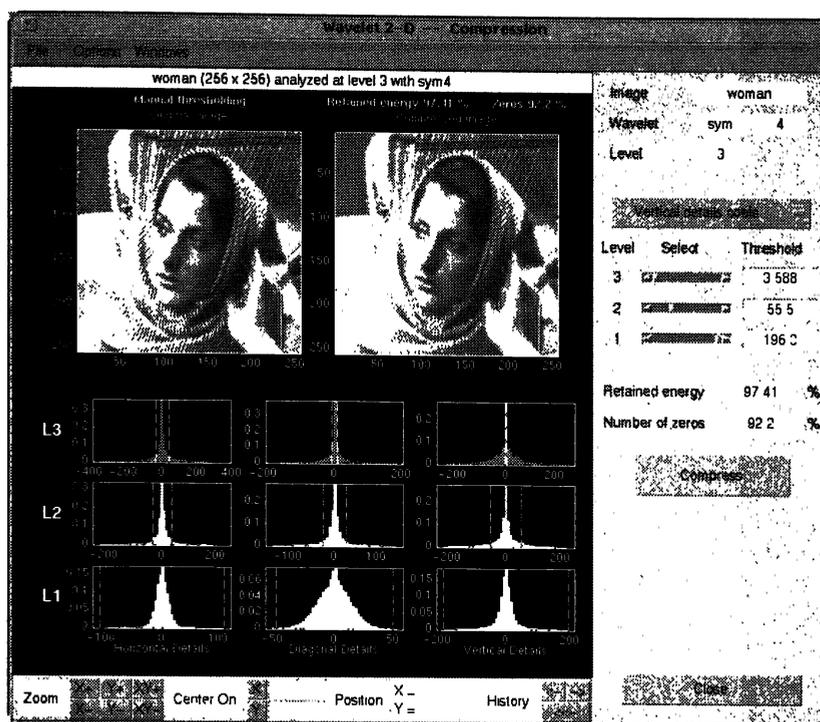


FIGURE 11 : Une image («Barbara») comprimée par ondelettes

juger de la qualité de la restitution, de nombreux critères numériques existent mais rien n'est plus imputoyable que l'œil humain.

La méthode est semblable à celle effectuant le débruitage. On conserve les coefficients de l'approximation la plus grossière auxquels on adjoint les coefficients de détail les plus forts, puis on reconstruit. Le réglage montré dans l'écran présenté par la figure 11 est manuel et s'effectue par niveau et par direction.

Examinons les histogrammes des niveaux de gris des coefficients de détail :

- en bas, figurent ceux concernant le niveau 1. Les droites verticales en pointillés contenant l'histogramme dans sa totalité, et ceci pour toutes les directions, tous les coefficients sont remplacés par zéro.
- pour le niveau 2, la stratégie est un peu moins sélective et conserve la majeure partie des coefficients grands en valeur absolue.
- enfin pour le niveau 3, critique pour la restitution des contours de l'image, seuls sont éliminés les très faibles coefficients.

L'image reconstruite en utilisant les coefficients seuillés est de très bonne qualité, en dépit du fort pourcentage de coefficients remplacés par zéro.

Ces performances peuvent être notablement améliorées en utilisant des ondelettes biorthogonales (une ondelette de décomposition avec suffisamment

de moments nuls afin que la représentation par ondelettes soit la plus creuse possible, associée à une ondelette duale de reconstruction très régulière et symétrique afin de supprimer le plus d'artefacts visuels) convenablement choisies. Nous en verrons une illustration plus loin.

X. UN SIGNAL COMPRIMÉ PAR ONDELETTES

Avant de passer à la compression d'empreintes digitales, mentionnons les méthodes de compression par ondelettes de signaux unidimensionnels, en traitant un exemple synthétique.

L'écran présenté dans la figure 12 est organisé en deux colonnes.

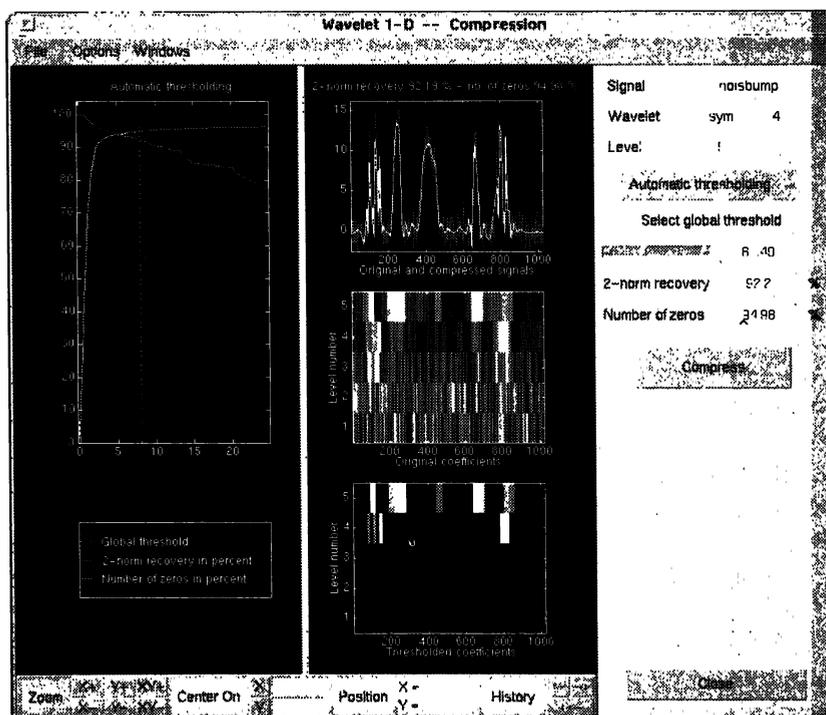


FIGURE 12 : Un signal comprimé par ondelettes

Dans la première colonne, on trouve un seul graphique permettant d'ajuster un seuil global pour les coefficients du signal à compresser : tous les coefficients qui sont, en valeur absolue, inférieurs au seuil (*cf.* la ligne verticale en pointillés) sont mis à zéro par la procédure de compression. En abscisse sont portées les valeurs possibles du seuil et en ordonnée deux courbes, l'une croissante donnant le pourcentage de zéros dans la représentation du signal comprimé, l'autre décroissante donnant le pourcentage de l'énergie conservée par le signal comprimé.

Dans la seconde colonne, le signal d'origine s est superposé au signal comprimé. En dessous on trouve une version coloriée des coefficients d'ondelettes des niveaux 1 à 5 du signal d'origine et dans le graphique encore en dessous, l'homologue pour les coefficients d'ondelettes seuillés, à partir desquels est reconstitué le signal comprimé.

Le pourcentage de zéros de la représentation par ondelettes, qui est un indicateur de la place occupée par le signal comprimé, est de 95 %, pour 92 % de l'énergie conservée (ceci vient du fait que le signal est bruité et donc que les fluctuations très rapides sont perdues). Les graphiques des coefficients avant et après l'opération de seuillage permettent de constater la rareté des coefficients conservés.

L'ondelette utilisée est sym4.

La compression de signaux unidimensionnels, bien que moins cruciale que la compression d'images, a de nombreuses applications parfois d'un grand intérêt économique. Par exemple pour les entreprises qui sont tenues de conserver avec une grande précision les profils de consommation individuelle de leurs clients, sur de longues périodes.

Comme le suggère la forme de l'interface proposée pour faire de la compression par un seuillage global, les idées sont très proches de celles du débruitage et l'algorithme mis en œuvre est identique. Il procède en trois phases dont le schéma de principe est le suivant :

- décomposition par l'ondelette ;
- on conserve les coefficients de l'approximation la plus grossière ainsi que les coefficients d'ondelettes les plus grands en valeur absolue ; les autres sont remplacés par zéro ;
- partant de ces coefficients modifiés, on reconstruit le signal comprimé.

XI. UNE EMPREINTE DIGITALE COMPRIMÉE PAR PAQUETS D'ONDELETTES

On trouve dans la figure 13 en haut à droite l'image originale et en bas à gauche l'image comprimée. En haut à gauche, on trouve l'arbre de paquets d'ondelettes et en bas à droite la décomposition de l'image à compresser. Cette décomposition est obtenue en décomposant non seulement les approximations en quatre (approximation et trois détails) mais en faisant aussi de même avec les détails. L'ondelette utilisée est celle de Haar, le résultat est bon mais peut être amélioré en utilisant une ondelette biorthogonale.

Dans la figure 14, on trouve en haut à gauche l'image originale qui est à compresser et à côté l'image comprimée. La méthode de compression procède par un seuillage global des coefficients de paquets d'ondelettes de l'image à compresser. Le graphique d'en dessous permet d'ajuster ce seuil global exactement comme en 1D. L'ondelette utilisée est une ondelette biorthogonale. Le résultat est très bon pour un pourcentage de zéros de la représentation de l'image comprimée de 95 %.

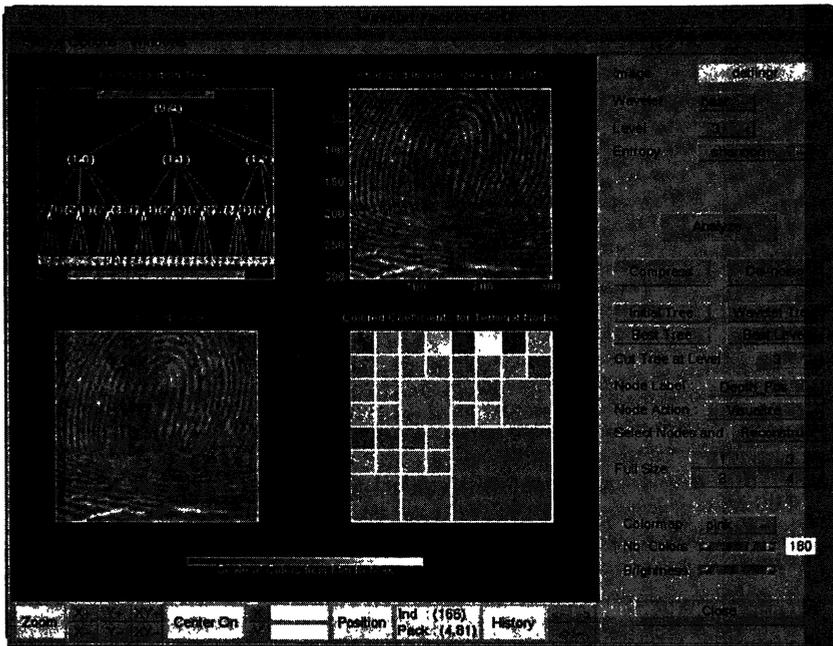


FIGURE 13 : Une empreinte digitale comprimée par paquets d'ondelettes

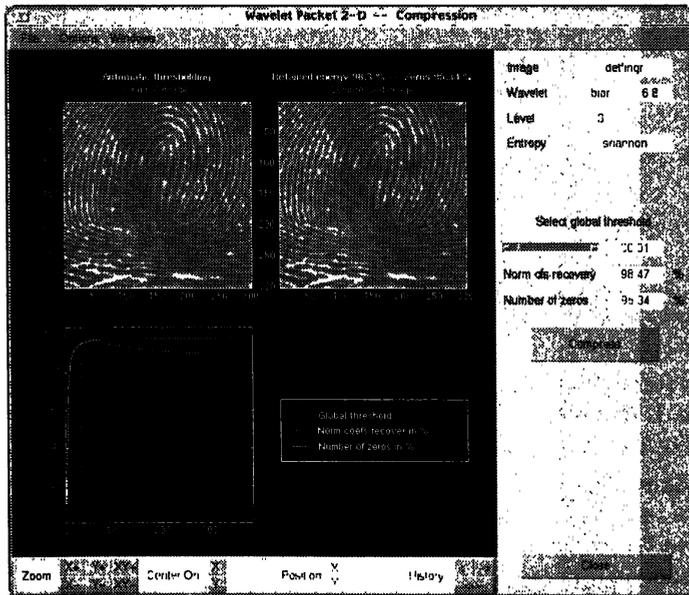


FIGURE 14 : Une empreinte digitale comprimée par paquets d'ondelettes

XII. QUELQUES MOTS DE CONCLUSION

Bien des questions n'apparaissent pas dans ce texte introductif, même parmi celles qui sont liées directement à la statistique. Passons donc en revue quelques thèmes complémentaires, sans chercher pour autant à être exhaustif.

- Estimation de paramètre. Une régression bien choisie sur l'échelle j de la variance $v(j)$ des coefficients d'ondelettes du niveau j , constitue un estimateur du paramètre H d'un processus Brownien fractionnaire, processus à temps continu défini sur $[0,1]$.
- Estimation fonctionnelle. De nombreux articles traitent en détail ce thème parfois assimilé au débruitage. Les ondelettes semblent avoir apporté un coup de fouet aux techniques d'estimation utilisant les bases de fonctions. Des théorèmes minimax sont produits et les espaces de Besov sont systématiquement explorés. On s'attaque à des fonctions nouvelles, localement assez irrégulières, pouvant présenter des motifs contrastés avec des zones régulières, des zones irrégulières, des variations locales rapides et des ruptures. Actuellement, l'estimation peut s'effectuer sans connaître la régularité globale de la fonction : elle s'adapte à la régularité. L'adaptation spatiale à la régularité locale est aussi un thème en développement.

Les ondelettes sont efficaces lorsque les fonctions cherchées se codent de façon parcimonieuse : peu de coefficients d'ondelettes suffisent à décrire chacune des fonctions à estimer.

Ces techniques servent à estimer des fonctions de nature différente : densité de probabilité, fonction de régression, fonction de hasard ou densité spectrale.

En outre, les estimateurs tirent avantage de l'extraordinaire rapidité algorithmique (linéaire en la taille des données) de la décomposition discrète.

La décomposition continue sert, entre autres, dans la caractérisation de l'irrégularité d'une fonction. Elle permet l'estimation du spectre de singularité d'une fonction f . Si $\alpha(x)$ est le degré de Hölder de f au point x , considérons l'ensemble des points x où le degré de Hölder vaut β . La dimension de Hausdorff $h(\beta)$ de cet ensemble quantifie sa taille.

La fonction h est appelée spectre de singularité de f . Son estimation est calculée à partir des lignes de maximum de la décomposition en ondelettes continue de f .

- Contrôle de qualité. Les ondelettes sont utilisées dans le domaine du contrôle de la qualité et la maîtrise statistique des procédés. Il s'agit essentiellement des aspects du diagnostic et de la surveillance, dans lesquels il s'agit de détecter des signaux courts, transitoires, de forme inhabituelle.
- Calcul scientifique. Les techniques d'ondelettes pénètrent l'analyse numérique où elles constituent les bases dans lesquelles les opérateurs s'approchent ou se représentent par des grandes matrices creuses. La turbulence pleinement développée constitue un des champs privilégiés d'application.

Mentionnons enfin qu'une source d'informations très complète sur les ondelettes et leurs applications est accessible sur le réseau international à l'adresse suivante : <http://www.amara.com/current/wavelet.html>.

BIBLIOGRAPHIE

- [1] ABRY P. (1997). *Ondelettes et turbulences*, Diderot.
- [2] ABRY P., SELLAN F., (1996). The wavelet-based synthesis for fractional Brownian motion proposed by F. Sellan and Y. Meyer : Remarks and fast implementation, *Applied and Comp. Harmonic An.*, 3, 377-383.
- [3] ANTONIADIS A., OPPENHEIM G., editors (1995). *Wavelets and statistics*, Lecture Notes in Statistics, 103, Springer Verlag.
- [4] ARNEODO A., ARGOUËL F., BACRY E., ELEZGARAY J., MUZY J.F. (1995). *Ondelettes, multifractales et turbulence*, Diderot ed.
- [5] BURKE-HUBBARD B. (1996). *The world according to wavelets*, A K Peters, Wellesley.
- [6] COHEN A. (1995). *Wavelets and multiscale signal processing*, Chapman and Hall.
- [7] COIFMAN R.R., WIRCKERHAUSER M.V. (1992). Entropy-based algorithms for best basis selection, *I.E.E.E Trans. on Inf. Theory*, vol. 38, 2, 713-718.
- [8] BRISLAWN C.M. (1995). Fingerprints go digital, *Notices of the AMS*, vol. 42, nov., 1278-1283.
- [9] BRUCE A., DONOHO D.L., GAO H-Y. (1996). Wavelet analysis, *I.E.E.E Spectrum*, oct., 26-35.
- [10] DAUBECHIES I. (1992). *Ten Lectures on Wavelets*, SIAM.
- [11] DE VORE R.A., JAWERTH B., LUCIER B.J. (1992). Image compression through wavelet transform coding, *I.E.E.E Trans. on Inf. Theory*, vol. 38, 2, 719-746.
- [12] DONOHO D.L. (1995). De-noising by soft thresholding, *I.E.E.E Trans. on Inf. Theory*, vol. 41, 3, 613-627.
- [13] DONOHO D.L., JOHNSTONE I.M. (1994). Ideal spatial adaptation by wavelet shrinkage, *Biometrika*, vol. 81, 425-455.
- [14] DONOHO D.L., JOHNSTONE I.M., KERKYACHARIAN G., PICARD D. (1995). Wavelet shrinkage : Asymptotia, *J.R.S.S.*, ser. B, vol 57, 2, 301-369.
- [15] HÄRDLE W., KERKYACHARIAN G., PICARD D., TSYBAKOV A. (1998). *Wavelets, approximation and statistical applications*, Lecture Notes in Statistics, 129, Springer Verlag.
- [16] KAHANE J.P., LEMARIÉ P.G. (1995). *Fourier series and wavelets*, Gordon and Research Publishers.
- [17] LEMARIÉ P. G., editor (1990). *Les Ondelettes en 1989*, Lecture Notes in Math., 1438, Springer Verlag.
- [18] MALLAT S. (1989). A theory for multiresolution signal decomposition : the wavelet representation, *I.E.E.E Trans. on P.A.M.I*, vol. 2, 7, 674-693.
- [19] MALLAT S. (1998). *A wavelet tour of signal processing*, Academic Press.
- [20] MEYER Y. (1990). *Ondelettes et Opérateurs*, Tome 1, Actualités mathématiques, Hermann.
- [21] MEYER Y. (1993). *Les Ondelettes. Algorithmes et Applications*, Colin ed., Paris, 2nde édition.
- [22] MISITI M., MISITI Y., OPPENHEIM G., POGGI J.M. (1994). Décomposition en ondelettes et méthodes comparatives : étude d'une courbe de charge électrique, *Revue de Statistique Appliquée*, vol. XLII, 2, 57-77.

MÉTHODES D'ONDELETTES EN STATISTIQUE

- [23] MISITI M., MISITI Y., OPPENHEIM G., POGGI, J.M. (1996). *Wavelet Toolbox User's Guide*, The Mathworks.
- [24] OGDEN T. (1997). *Essential wavelets for statistical applications and data analysis*, Birkhauser.
- [25] SELLAN F. (1995). Synthèse de mouvements browniens fractionnaires à l'aide de la transformation par ondelettes, *C.R.A.S*, 321, 351-358.
- [26] TORRÉSANI B. (1995). *Analyse continue par ondelettes*, CNRS Ed.
- [27] WIRCKERHAUSER M.V. (1994). *Adapted Wavelet Analysis from Theory to Software Algorithms*, AK Peters.